

# Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/140543/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Guo, Pengfei, He, Lang, Liu, Shuangyin ORCID: <https://orcid.org/0000-0003-4544-3481>, Zeng, Delu and Liu, Hantao ORCID: <https://orcid.org/0000-0003-4544-3481> 2022. Underwater image quality assessment: subjective and objective methods. IEEE Transactions on Multimedia 24 , pp. 1980-1989. 10.1109/TMM.2021.3074825 file

Publishers page: <http://dx.doi.org/10.1109/TMM.2021.3074825>  
<<http://dx.doi.org/10.1109/TMM.2021.3074825>>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies.

See

<http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



# Underwater Image Quality Assessment: Subjective and Objective Methods

Pengfei Guo, Lang He, Shuangyin Liu, Delu Zeng, Hantao Liu

Underwater image enhancement plays a critical role in marine industry. Various algorithms are applied to enhance underwater images, but their performance in terms of perceptual quality has been little studied. In this paper, we investigate five popular enhancement algorithms and their output image quality. To this end, we have created a benchmark, including images enhanced by different algorithms and ground truth image quality obtained by human perception experiments. We statistically analyse the impact of various enhancement algorithms on the perceived quality of underwater images. Also, the visual quality provided by these algorithms is evaluated objectively, aiming to inform the development of objective metrics for automatic assessment of the quality for underwater image enhancement. The image quality benchmark and its objective metric are made publicly available.

**Index Terms**—Underwater image, image quality assessment, perception experiment, statistical analysis, objective metric.

## I. INTRODUCTION

**A**QUACULTURE is an important component of the marine industry. In order to reduce the costs of human resources and protect the marine environment, research has been undertaken to develop underwater robots for aquaculture [1]. The main task of underwater robots is to detect and track underwater objects based on capturing underwater images. To make these images useful for underwater tasks, image enhancement is often applied. However, the complex underwater environment makes the enhancement of underwater images a very challenging problem. Firstly, the underwater images exhibit the haze-like property which is caused by lights scattered and deflected by plankton, sand and minerals in the ocean [2]. Secondly, the underwater images present various color changes resulted from varying degrees of light

attenuation for different wavelengths [3]. Moreover, with the influence of the artificial lights and unstable conditions for capturing raw images, the noise is hard to be expressed by some known statistical distributions in the underwater images [4].

For underwater image enhancement, the designed algorithms should solve the following issues: (a) the color-casting (greenish/bluish color) problem caused by light attenuation. (b) the sharp-degradation (blurring effects) problem due to forward scattering of light, especially, by enhancing the edges and details of images will improve further processing such as object detection, segmentation, saliency detection and so on. (c) the contrast degradation (hazing) resulted from backward scattering of light. However, the special properties of the underwater images make the direct use of traditional image enhancement methods rather infeasible [5].

Many dedicated algorithms have been proposed to enhance the quality of underwater images. These algorithms can be categorized into the physics-based approach and the non-physics-based approach. The physics-based methods model the processes of underwater optical imaging. This can be described by the image formation model (IFM) [6], [7], [8]:

$$I^c(i, j) = J^c(i, j)t^c(i, j) + B^c(1 - t^c(i, j)), c \in \{r, g, b\}, \quad (1)$$

where  $I^c(i, j)$  is the intensity of the scene captured by the camera in the  $c$ -channel at the position  $(i, j)$ ,  $J^c(i, j)$  is the restored underwater scene in the  $c$ -channel at the position  $(i, j)$ ,  $B^c$  is the homogeneous background light in the  $c$ -channel, and  $t^c(i, j)$  is the transmission medium map in the  $c$ -channel at the position  $(i, j)$ . The enhancement of underwater images using the IFM can be also regarded as the IFM-based image restoration methods, e.g., dark channel prior (DCP) [9], underwater dark channel prior (UDCP) [10], minimum information loss prior (MIP) [11], generalized dark channel prior (GDCP) [12], non-local image dehazing (NLD) [13], restoration based on image blurriness and light absorption (IBLP) [14] and so on. The methods that do not adopt the information of underwater optical imaging are called the non-physics-based methods and can be regarded as the IFM-free methods. The IFM-free based methods mainly deal with the improvement of color, edges and contrast of underwater images based on the pixel intensity re-distribution. The IFM-free methods range from the traditional techniques (such as contrast limited adaptive histogram equalization (CLAHE) [15], decorrelation stretch (DS) [16], image fusion (IF) [17] and so on) to data-driven techniques (e.g., generative adversarial network (GAN) [18] and convolutional neural network (CNN) [19]).

P.Guo is with the school of Computational Science, Zhongkai University of Agriculture and Engineering, Guangzhou, 510225 China, and also with the school of Mathematics, South China University of Technology, Guangzhou, 510641, China (e-mails: Guopfzhku@163.com).

L.He is with Shool of Computer Science and Engineering, Sun Yat-sen University, Guangzhou 510006, China (e-mails: helang5@mail2.sysu.edu.cn).

S.Liu is with College of Information Science and Technology, Zhongkai University of Agriculture and Engineering, Guangzhou, 510225, China (e-mails: SandyGuo480981@outlook.com.com).

D. Zeng is with the school of Mathematics, South China University of Technology, Guangzhou, 510641, China (e-mail: dlzeng@scut.edu.cn).

H. Liu is with the School of Computer Science and Informatics, Cardiff University, Cardiff, CF24 3AA United Kingdom (e-mail: LiuH35@cardiff.ac.uk)(corresponding author).

This work is funded by Foundation for High-level Talents in Higher Education of Guangdong Province under Grant (2017KTSCX095, 2016KQNCX075), the general program of technology project of Guangzhou (201707010221), the Lifting project of Guangzhou Youth Talent, and Guangdong basic and applied basic research foundation (2020A1515110958).

Notwithstanding the progress made on the development of the image enhancement algorithms for underwater images, there is little work on the comparative study of these algorithms and a lack of a generic framework for the quality assessment of the enhanced images.

In this paper, we selected five state-of-the-state underwater image enhancement algorithms including both IFM-free methods and IFM-based methods. Firstly, We describe and compare these underwater enhancement algorithms and analyse the signal properties of these algorithms. Secondly, we build a benchmark of underwater image enhancement through subjective image quality assessment. Furthermore, we propose a novel objective quality metric for underwater image enhancement that can automatically assess the quality of enhanced images (note quality assessment must be reliant on the enhanced image only; there is no reference of perfect/maximum quality as an enhanced image may be of higher or lower quality relative to the original).

The contributions of this paper are:

- A comparative study of state-of-the-art algorithms for underwater image enhancement;
- A benchmark for the assessment of underwater image enhancement algorithms;
- An improved objective metric for underwater image quality assessment.

The rest of the paper is organized as follows: Section II evaluates different underwater image enhancement algorithms. Section III presents our new underwater image enhancement benchmark and the subjective and objective evaluation of the quality of the enhanced images. Section IV concludes the paper.

## II. STATE-OF-THE-ART UNDERWATER IMAGE ENHANCEMENT ALGORITHMS

### A. Underwater image enhancement algorithms

Recently, many image processing methods are adopted to study the underwater image enhancement problem. Traditionally, Histogram Equalization (HE) [20] and Modified Histogram Equalization (MHE), e.g., Contrast Limited Adaptive Histogram Equalization (CLAHE), and wavelet-HE are considered as a typical tool that can be used to improve global contrast of the low-light images. Qiao et. al. [21] presented a CLAHE-wavelet enhancement algorithm for the underwater sea cucumber images. They use CLAHE to increase the contrast of underwater images with Rayleigh prior distribution. The enhanced image is further improved by de-noising using soft threshold wavelet transform. Ahmad et.al. [22] proposed a dual-image Rayleigh-stretched contrast-limited adaptive histogram specification algorithm to combine local and global contrast and use color correction to improve the visual quality of underwater images.

To deal with the color casting problem of underwater images, the classical color correction methods such as White Balance (WB) and Gray-Edge Assumption (GEA) are used to adaptive change the saturation of the images to upgrade their visual quality. These traditional color correction algorithms are used as the pre-processing or post-processing procedures of

underwater enhancement algorithms [23]. Due to the influence of artificial light and low-light environment, the use of contrast enhancement algorithms and color correction algorithms often give over-enhancement effects such as serious artificial details, and halos.

Image fusion is another successful strategy for IFM-free underwater enhancement. In 2012, Ancuti et.al. reported an image fusion-based underwater enhancement algorithm [24]. The fused images are produced by using white balance to correct color and using local adaptive histogram equalization to increase contrast. The enhanced images with more detailed information and better contrast are then obtained by using underwater associated adaptive weighted multi-scale method. The follow-up work of Ancuti et.al. presented a novel image fusion method for underwater images based on color balance. The fused images are obtained by using Gamma correction and sharpness algorithms to solve the white balance problem. The fused images deal with the low-light disadvantage of underwater images, and show better quality in the dark regions of underwater images [17].

Some researchers study the underwater image enhancement problem based on the image formation model. Yang et.al. [10] proposed a modified dark channel prior algorithm to enhance the low complexity underwater images. They estimated the depth map by median filter instead of the soft matting. The restoration images are further improved by Gamma correction to increase the contrast of the images. Li et. al. [11] proposed a two-step underwater image enhancement method. The first step employed a dehazing algorithm with minimum information loss that restores the natural images. The second step was to increase the local and global contrast of underwater images by histogram distribution prior. Peng et. al. [14] presented an underwater image restoration algorithm based on light absorption and image blurriness. Based on estimation of the background light from blurry regions of underwater images, the transmission map and depth map gave a good enhanced underwater image. Peng et. al. [12] proposed a generalized dark channel prior algorithm (GDCP) for underwater images. They employed the scene transmission map estimated by scene ambient light differential instead of the transmission map. This proposed GDCP was the first method to study the effect of the color-dependent color change for degraded images caused by light absorption and scattering. Dana et.al. [13] reported a non-local image dehazing algorithm based on non-local prior instead of patch-based prior. The performance of IMF-based algorithms depends on the accuracy of the estimated depth map and transmission map of underwater images.

In summary, the IFM-based methods use the image formation model to transform the enhancement problem into an inverse problem which can be solved by estimating the background light map and transmission medium map. The IFM-free methods attempt to improve the image quality by dynamically changing the histogram of the pixel intensities. It is also worth noting that recent studies demonstrate the potential of the deep learning-based underwater image enhancement algorithms [25], [26], [27], [28]. However, due to the lack of sufficient and high-quality training data, these algorithms often rely on synthetic images for model training, therefore,



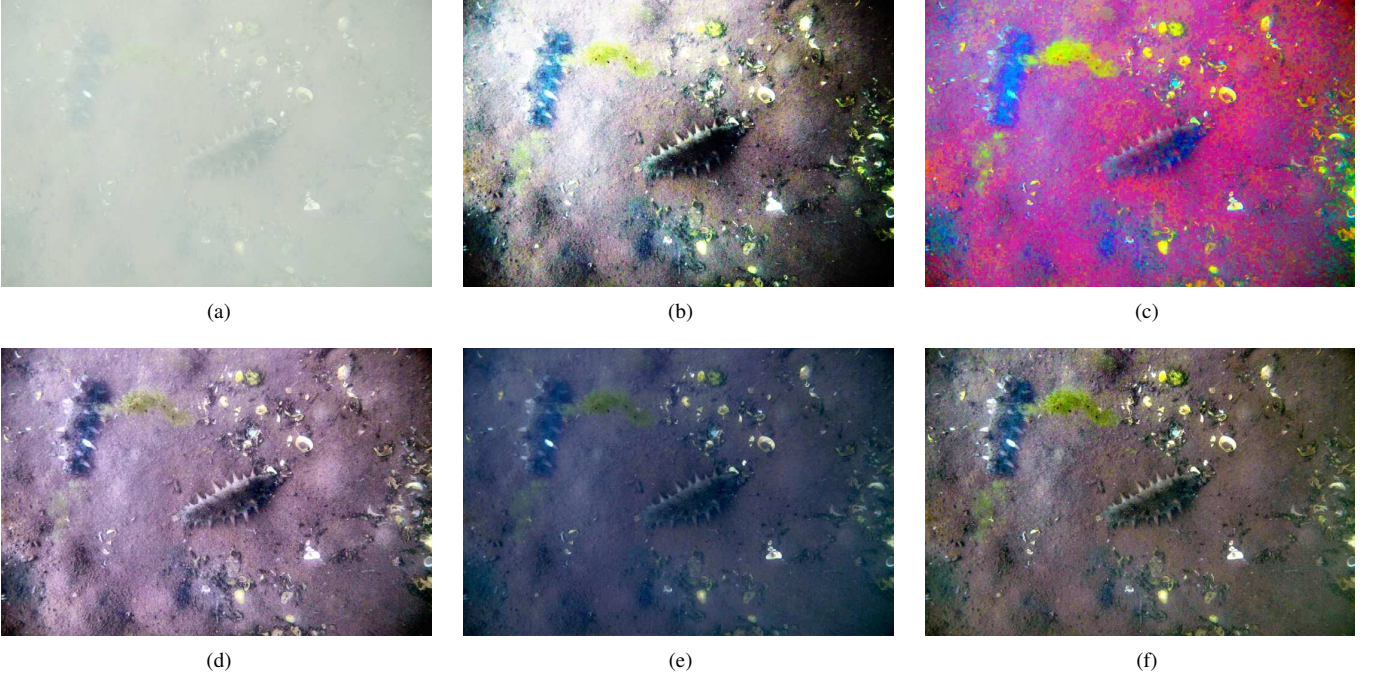


Fig. 1. Illustration of the effect of state-of-the-art enhancement algorithms applied to underwater images. (a) original image, (b)-(f). images enhanced by HE [15], DS [16], IF [17], UDCP [10], and NLD [13] algorithm, respectively.

are difficult to generalise to the entire space of real underwater scenes.

We describe below the detail of the state-of-the-art underwater image enhancement algorithms, focusing on IFM-based and IFM-free methods.

### 1) Contrast Limited Adaptive Histogram Equalization (HE)

The histogram equalization (HE) algorithm has a good performance for images of relatively equalized intensities. Since the histogram of underwater images concentrates on some ranges of intensities, we chose a modified histogram equalization algorithm called Contrast Limited Adaptive Histogram Equalization (CLAHE). The CLAHE method processes an image patch by patch. For each patch, an adaptive threshold is selected to control the noise caused by local histogram equalization. The result of CLAHE for an RGB image is obtained by using the CLAHE algorithms on each channel [15].

### 2) Decorrelation Stretch Enhancement (DS)

Decorrelation stretch is a PCA-based linear and pixel-wise transformation conducted by the values of original and enhanced image statistics [16]. For an RGB image  $I$ , set input channel image  $I_c = [\mathbf{x}_1, \dots, \mathbf{x}_h]$  and target channel image  $Z_c = [\mathbf{z}_1, \dots, \mathbf{z}_h]$ . The covariance matrix of  $I_c$  is  $\Sigma_{I_c} = I_c^T I_c$ , and the mean vector of each column of  $I_c$  is  $\mathbf{m} = \frac{1}{h} \sum_{k=1}^h \mathbf{x}_k$ . The singular value decomposition (SVD) of  $\Sigma_{I_c}$  can be expressed as  $\Sigma_{I_c} = V_c^T \Lambda V_c$ . By defining a diagonal matrix  $S_c = \text{diag}\{\Sigma_{I_c}(k, k)\}_{k=1, \dots, h}$ , and  $H_c = \text{diag}\{\frac{1}{\sqrt{\Lambda(k, k)}}\}_{k=1, \dots, h}$ , the model of decorrelation stretch can be expressed as:

$$\mathbf{z}_i = S_c V_c H_c V_c^T (\mathbf{x}_i - \mathbf{m}) + \mathbf{m}. \quad (2)$$

The decorrelation stretch can improve the image quality by increasing the color differences.

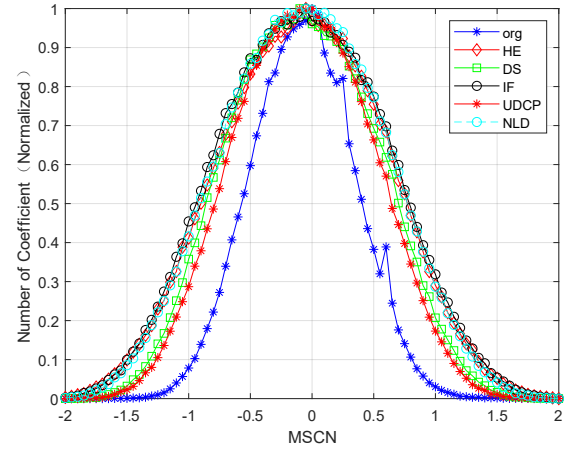


Fig. 2. Illustration of the histogram of MSCN coefficients for the enhanced images (in comparison to the original image) of Fig. 1. In the figure, it shows the MSCN histograms of the HE enhanced image (Fig.1.(b)), the DS enhanced image (Fig.1.(c)), the IF enhanced image (Fig.1.(d)), the UDCP enhanced image (Fig.1.(e)), the NLD enhanced image (Fig.1.(f)) and the original image (Fig.1.(a)).

### 3) Image Fusion based Underwater Image Enhancement (IF)

The image fusion enhancement algorithm operates on each channel of RGB [17]. The enhanced fusion image can be computed by the following formula:

$$Z_c(p, q) = \sum_{k=1}^r \bar{W}_c^k(p, q) I_c^k(p, q), c \in \{r, g, b\}, \quad (3)$$

where  $I_c^k$  is the  $k$ -th input image in the  $c$  channel,  $\bar{W}_c^k$  is the  $k$ -th normalized weight map in the  $c$  channel ( $\sum_{k=1}^r \bar{W}_c^k = 1$ ).

for any  $(p, q)$ , and  $(p, q)$  is the position in image  $I_c^k$  and  $Z_c$ .

The above naive fusion strategy may bring in some undesirable artifacts and halos. The Laplacian multi-scale image fusion method is adopted to overcome the disadvantages of the direct fusion method. In this multi-scale image fusion algorithm, let  $G_i$  denote a sequence of operators including low-pass Gaussian filtering, decimation, and up-sampling operations. The Laplacian pyramid decomposition form of image  $I_c$  is:

$$I_c(p, q) = \sum_{i=1}^L L_i(p, q) = \sum_{i=1}^L [G_{i-1}(p, q) - G_i(p, q)], \quad (4)$$

where  $L_i(p, q)$  is the  $i$ -th level of Laplacian pyramid,  $G_i(p, q) = G_i(I_c(p, q))$  is the  $i$ -th level of Gaussian pyramid and  $G_0(p, q) = I_c(p, q)$ . Then the Laplacian pyramid based multi-scale image fusion model can be expressed as:

$$R_{c,i}(p, q) = \sum_{k=1}^r G_i[\bar{W}_c^k(p, q)] L_i[I_c^k(p, q)], \quad (5)$$

where  $R_{c,i}$  is the  $i$ -level fused image, and  $G_i[\bar{W}_c^k]$  is the  $i$ -level Gaussian pyramid of the normalized weight map (note there are good weight map choices taking into account the characteristics of underwater images [17], [24], [29]).

#### 4) Underwater Dark Channel Prior Enhancement (UDCP)

The underwater dark channel prior algorithm is a IFM-based enhancement method [10]. We rewrite the information formation model as:

$$\mathbf{I}(\mathbf{x}) = \mathbf{J}(\mathbf{x})\mathbf{t}(\mathbf{x}) + \mathbf{B}(1 - \mathbf{t}(\mathbf{x})), \quad (6)$$

where  $\mathbf{x}$  is the pixel position of an image,  $\mathbf{I}$  is the observed image,  $\mathbf{J}$  is the enhanced image,  $\mathbf{t}$  is the transmission medium map, and  $\mathbf{B}$  is the background light.

The dark channel of the enhanced image (haze-free image)  $\mathbf{J}$  can be defined as:

$$J_{dark}(\mathbf{x}) = \min_{c \in \{r, g, b\}} (\min_{\mathbf{y} \in \Omega(\mathbf{x})} J_c(\mathbf{y})), \quad (7)$$

where  $J_c$  is the  $c$  channel of image  $\mathbf{J}$ ,  $\Omega(\mathbf{x})$  is the neighborhood of  $\mathbf{x}$ . It is well-known that the haze-free image should satisfy the following prior:  $J_{dark}(\mathbf{x}) = 0$ .

The background light is estimated by selecting the brightest position in the dark channel:

$$\tilde{B}_c(I) = I_c(\arg \max_{\mathbf{x} \in \mathbf{I}} I_{dark}(\mathbf{x})), \quad c \in \{r, g, b\}, \quad (8)$$

In a local patch  $\Omega(\mathbf{x})$ , the transmission medium map  $\tilde{t}(\mathbf{x})$  is assumed as a constant. The transmission medium map  $\tilde{t}(\mathbf{x})$  can be obtained by the operation of minimizing the equation (6) in the local patch for three channels of the image based on haze-free image prior:

$$\tilde{t}(\mathbf{x}) = 1 - \min_{c \in \{r, g, b\}} (\min_{\mathbf{y} \in \Omega(\mathbf{x})} (I_c(\mathbf{y})/\tilde{B}_c)). \quad (9)$$

Combining the equations (8) and (9) with equation (6), the enhanced image  $\mathbf{J}$  can be estimated by the following model:

$$\tilde{\mathbf{J}}(\mathbf{x}) = \frac{\mathbf{I}(\mathbf{x}) - \tilde{\mathbf{B}}}{\max(\tilde{t}(\mathbf{x}), t_0)} + \tilde{\mathbf{B}}, \quad (10)$$

where 0.1 is one of the typical values of  $t_0$ .

#### 5) Non-local Dehazing Enhancement (NLD)

The input image of the non-local dehazing algorithm is regarded as a 3D RGB image [13]. The model of the non-local dehazing method is the same as equation (6). Define a new image  $\mathbf{I}_B$  as:

$$\mathbf{I}_B(\mathbf{x}) = \mathbf{I}(\mathbf{x}) - \tilde{\mathbf{B}}, \quad (11)$$

where  $\tilde{\mathbf{B}}$  is the background light estimated by equation (8).  $\mathbf{I}_B(\mathbf{x})$  can be rewritten as the spherical coordinate:

$$\mathbf{I}_B(\mathbf{x}) = [\mathbf{r}(\mathbf{x}), \theta(\mathbf{x}), \phi(\mathbf{x})]. \quad (12)$$

The pixel-wise transmission medium map based on the non-local haze-free prior can be estimated as:

$$\tilde{t}_{NLD}(\mathbf{x}) = \frac{r(\mathbf{x})}{\hat{r}_{max}(\mathbf{x})}. \quad (13)$$

where  $\hat{r}_{max}(\mathbf{x}) = \max_{\mathbf{x} \in H} (r(\mathbf{x}))$ , and  $H = \{\mathbf{x} \in \mathbf{I} | \theta(\mathbf{x}) = \phi(\mathbf{x})\}$  is a haze line. In order to control the background light of the dehazing image, the final estimated transmission medium map is given by:  $\tilde{t}_{LB}(\mathbf{x}) = \max\{\tilde{t}_{NLD}(\mathbf{x}), t_{LB}(\mathbf{x})\}$ , where  $t_{LB}(\mathbf{x}) = 1 - \min_{c \in \{r, g, b\}} (I_c(\mathbf{x})/\tilde{B}_c)$  is a low bound of transmission medium map.

The non-local dehazing enhanced image  $\tilde{\mathbf{J}}_{NLD}$  can be computed as:

$$\tilde{\mathbf{J}}_{NLD}(\mathbf{x}) = \frac{\mathbf{I}(\mathbf{x}) - \tilde{\mathbf{B}}}{\tilde{t}_{LB}(\mathbf{x})} + \tilde{\mathbf{B}}. \quad (14)$$

#### B. Impact of enhancement algorithms on image quality related signal properties

The above selected algorithms reflect diverse approaches in image enhancement, including three IFM-free algorithms (HE, DS, and IF) and two IFM-based algorithms (UDCP and NLD). For the IFM-free approach, HE, DS and IF re-distribute the histogram of image intensities with different representation models and their corresponding natural scene statistical priors. More specifically, the HE algorithm alters the intensity distribution in the single colour model space (RGB) based on the grey-level assumption; the DS algorithm produces a large range of image intensities using the high dimensional colour space centering prior by considering the RGB colour model as a 3D vector space; and the IF algorithm wraps the advantages of preprocessing enhancement algorithms and reduces the artificial noise by the Laplacian image representation model. For the IFM-based approach, the UDCP and NLD algorithms attempt to compensate the degradation of light across water by solving the IFM inverse problem using different physical priors. More specifically, the UDCP algorithm tackles the inverse problem considering an RGB image as three independent channels; while the NLD algorithms considers an RGB image as a 3D vector space which corresponds to the high-level physical prior. Overall, all algorithms have a significant effect on the pixel-wise intensity. Fig. 1 shows the effect of applying these enhancement algorithms on an underwater image. Now, this paper will investigate the output image quality of these algorithms.

The histogram of mean subtracted contrast normalized (MSCN) coefficients [30] represents important statistical properties of an image. Based on the circular-symmetric Gaussian function sampling, the MSCN coefficients can be computed by the local estimates of the weighted mean  $\mu(\mathbf{x})$  and the weighted variance  $\sigma(\mathbf{x})$  in a neighborhood of  $\mathbf{x}$ , and the MSCN coefficients at the image position  $\mathbf{x}$  can be expressed as [30]:

$$MSCN(\mathbf{x}) = \frac{I(\mathbf{x}) - \mu(\mathbf{x})}{\sigma(\mathbf{x}) + \epsilon}, \quad (15)$$

where  $\epsilon$  is a small positive constant. For natural scene images, Ruderman et. al. [31] first found that the histogram of MSCN coefficients follow a Gaussian distribution. The above equation means that the local irregularities of the image can be represented by the histogram of MSCN coefficients. Adrian et. al. [30] reported that the change of visual quality of an image is directly proportional to the variance of the histogram of MSCN coefficients. We compare the histogram of MSCN coefficients of an original image and its enhanced versions. Fig.2 shows the histogram of MSCN coefficients for the images of Fig. 1. It can be seen that the enhancement algorithms enlarge the variances of the histograms which implies the increase in image quality.

### III. PROPOSED METHODS FOR SUBJECTIVE AND OBJECTIVE QUALITY ASSESSMENT

#### A. Image acquisition

To form a framework for the assessment of underwater image enhancement algorithms, we acquired a set of new images in a sea cucumber farm at Qingdao Ocean Lake in Shandong Province, China. The sea cucumber fishery stays uncovered with the depth of water about 3-4m. The image capturing equipment is a C-Watch remotely operated underwater vehicle (ROV) which is composed of a battery pack and a digital camera. The battery pack can provide power to the electrics and propulsion units which include a pressure sensor, GPS sensor, attitude and heading reference unit, and camera with a housing device. The model of camera is Canon PowerShot G12. The camera faces down and focuses on the sea cucumbers which live in different complex backgrounds in a shallow pool. The resolution of the images with JPEG format is  $1280 \times 720$  pixels. The captured underwater images have the following characteristics: (1) the distance between sea surface and object is about 3-4m; (2) the brightness of the images is caused by sunlight (no artificial light); and (3) most of the photos are taken from top to bottom. Fig. 1(a) illustrates an exemplar image captured by the equipment.

#### B. Database of underwater image enhancement

In order to analyse the output image quality of the underwater enhancement algorithms, an image quality assessment database needs to be constructed. Considering the inherent limitations of performing a controlled perception experiment, a set of source images reflecting the diversity in image content is often selected, e.g., the LIVE database selected 29 source images to reflect the space of diverse natural scenes[32]. We selected 40 source images from over 1800 captured images

based on the following criteria: the selected images should reflect scene complexity combined with various underwater environments, including color-casting, sharp-degradation, and contrast degradation problems. We applied the above mentioned five enhancement algorithms, HE, DS, IF, UDCP, and NLD on the 40 selected images. By doing so, we created a benchmark dataset of underwater image enhancement which consists of 40 original images and 200 enhanced images. We now perform experiments to evaluate the enhanced images produced by these algorithms as follows.

#### C. Subjective quality assessment

##### 1) Perception experiment

A fully-controlled perception experiment was conducted to obtain a benchmark (i.e., ground truth) for the quality of underwater image enhancement. The experiment followed the guidelines described in [33], [34]. A single-stimulus method was used, in which human subjects were asked to score the overall quality for each stimulus in the absence of any reference. The scoring scale ranged from 0 to 100, and included additional semantic labels, i.e., Bad, Poor, Fair, Good and Excellent at intermediate points to help subjects with expressing their opinions on the numerical scale.

The subjective experiments were carried out in a standard office environment with approximately constant ambient light. The test stimuli were displayed on a 19-inch LCD screen, with a native resolution of  $1920 \times 1080$  pixels. The viewing distance was maintained around 60cm. The participants of the study consisted of eighteen viewers naive to image quality assessment. They were nine males and nine females and between the ages of 23 and 52. Before the start of each experiment, written instructions describing the experimental procedure, including the scoring scale and timing, were given to each subject. A training set of five typical images was presented to the participants in order to familiarise them with the issues in underwater image quality and with use of the scoring scale. The stimuli used in training were different from those used for the real assessment. After training, the test images were presented to each participant in a random order. Each stimulus was shown once, and the participants were requested to assess its quality immediately after viewing it. To avoid fatigue, each session per subject was divided into two sub-sessions of having half amount of stimuli each and with a 10-minute break between sub-sessions.

##### 2) Mean opinion scores

First, z-scores were calculated to account for the differences between subjects in the use of the scoring scale and calibrate them towards the same mean and standard deviation [35], [36]. The raw subjective scores were converted into z-scores as follows:

$$z_{ij} = \frac{r_{ij} - \mu_i}{\sigma_i}, \quad (16)$$

where  $r_{ij}$  denotes the raw score given by the  $i$ -th subject to the  $j$ -th test stimulus,  $\mu_i$  is the mean of all scores for the subject  $i$ , and  $\sigma_i$  is the corresponding standard deviation.

Second, an outlier detection and subject removal procedure as suggested in [35] was performed, where results more than

two standard deviations from the mean for a test image were considered to be outliers and an individual was an outlier if 20% of scores submitted were outliers. This procedure resulted in 4% of scores being detected and removed as outliers (note the outliers are evenly distributed across different enhancement algorithms), and no participant being rejected. Finally, the mean opinion score (MOS) of each stimulus was computed as the mean of the remaining z-scores over all subjects:

$$MOS_j = \frac{1}{s} \sum_{i=1}^s z_{ij}, \quad (17)$$

where  $s$  is the number of remaining scores (after outlier removal) for the  $j$ -th image. To make the final scores easier to interpret, the resulting mean opinion scores were linearly remapped to the range of [0, 10], as the histogram shown in Fig. 3. This gives a new image quality assessment database for underwater image enhancement (UEIQA).

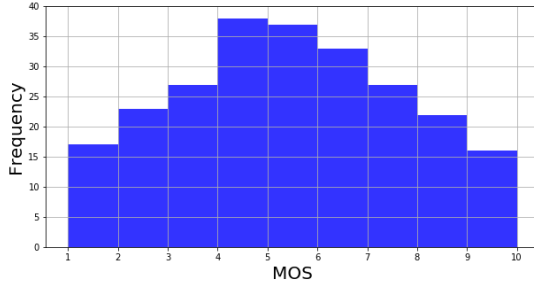


Fig. 3. Illustration of the histogram of subjective ratings (MOS) of stimuli contained in the UEIQA database.

### 3) Statistical properties of the UEIQA database

Since the ground truth image quality (i.e., MOS) is obtained, this can be used to assess the performance of underwater image enhancement algorithms. Fig. 4 illustrates the impact of the five different algorithms on image quality, averaged over all corresponding test images. The observed tendencies were statistically analysed with a hypothesis testing. Pairwise comparisons were performed with a t-test preceded by a test for the assumption of normality (note in the case of non-normality, a non-parametric version analogue to the t-test is applied), selecting the enhancement algorithm as the independent variable and the image quality as the dependent variable. The test reveals the following order in quality (Note "<" means is statistically significantly ( $p$ -value < 0.05) less than, except for the case that jointly underlined algorithms do not significantly differ):  $DS < Original < HE < UDCP < \underline{NLD} < \underline{IF}$ . The ground truth shows that using algorithms HE, UDCP, NLD and IF improves image quality relative to original images (i.e., the quality of enhanced images are statistically significantly better than that of the original images), and that NLD and IF outperforms other algorithms in term of final image quality. IF and NLD algorithms (see Fig. 1. (d) and (f)) successfully remove haze from the original images while preserving the details of the main objects. Also, the results show that the DS algorithms deteriorates the quality of original images significantly, which means DS fails in

enhancing underwater images. This might be due to the fact that the DS algorithm yields significant changes in colour to the original images (see Fig. 1. (c)), and consequently enhances unimportant objects in the underwater scenes.

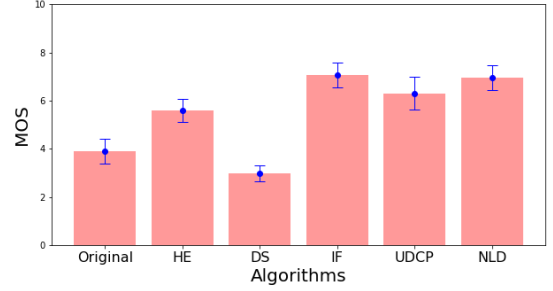


Fig. 4. Mean opinion score (MOS) averaged over all corresponding test images for each algorithm in our subjective quality experiment. The error bar indicates a 95% confidence interval.

### D. Objective evaluation of visual quality

Subjective experiments when properly conducted are the most reliable means of assessing image quality. However, subjective testing is time-consuming, expensive, and impractical in many circumstances. A more realistic way to evaluate the output quality of enhancement algorithms is to use computational models, which can automatically predict the quality of images as perceived by humans. These objective metrics can be classified into full-reference (FR) ([37], [38]), reduced-reference (RR) ([39], [40], [41]), and no-reference (NR) ([42], [43], [44], [45], [46], [47], [48]) approaches, depending on whether the original image is used as the reference of highest quality [37], [49], [50]. But in image enhancement the original image does not represent such reference of perfect/maximum quality as an enhanced image may be of higher or lower quality relative to the original. Thus, our objective evaluation for underwater image enhancement must rely on no-reference approaches.

Now, the first attempt is to investigate whether existing image quality metrics can assess the quality of enhanced images. We used six state of the art metrics, including three general-purpose NR image quality metrics (i.e., Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE) [51], Naturalness Image Quality Evaluator (NIQE) [52], and Blind Image Quality Assessment in the DCT Domain (BLIIND-II) [53]) and two dedicated underwater image quality metrics (i.e., Underwater Color Image Quality Evaluator (UCIQE) [54], and Underwater Image Quality Metric (UIQM) [55]). Each metric was applied to assess the quality of the images contained in our new UEIQA database, resulting in an objective image quality score (IQS) per image. The Person linear correlation (CC) between the subjective MOS and objective IQS is calculated and is 0.4148, 0.3999, 0.1718, 0.5230, 0.6159 for BRISQUE, NIQE, BLIIND-II, UCIQE, and UIQM, respectively (note the detail of the performance evaluation will be given below). BRISQUE, NIQE, BLIIND-II metrics were mainly devised to assess artifacts in image compression and transmission,



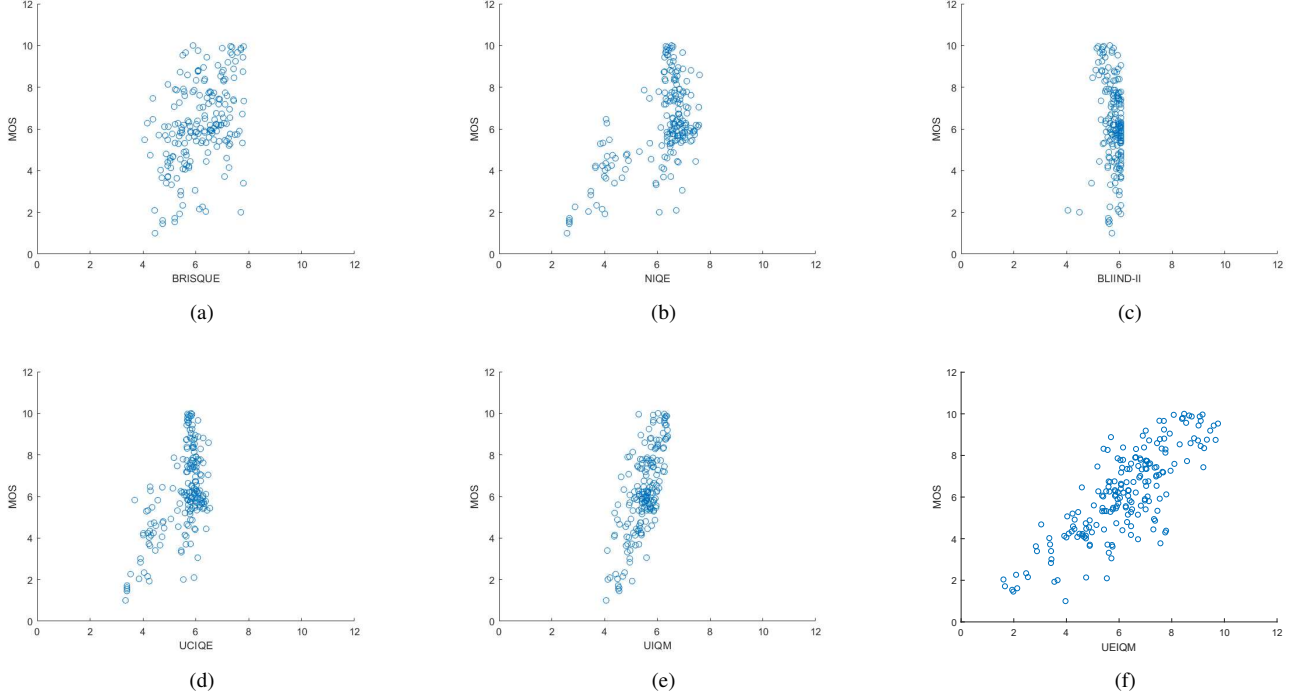


Fig. 5. Scatter plot of MOS versus BRISQUE, NIQE, BLIND-II, UCIQE, UIQM, and proposed UEIQM, respectively.

and their CC values are rather low. UCIQE, and UIQM metrics give better performance, but unfortunately, they are not adequate for assessing the quality of underwater image enhancement. Hence, our next attempt is to build a problem-specific objective metric. Note, the goal here is to investigate a proof-of-concept model, and developing a sophisticated quality metric for underwater image enhancement will be treated in a separate contribution in our future research.

#### 1) Proposed Metric

We extracted the underwater enhancement related image features including underwater image colorfulness measure, underwater image sharpness measure, and underwater image contrast measure proposed by Panetta et. al. [55].

With the influence of light attenuation, underwater images often display a greenish and bluish appearance. The low-light condition of the environment is the other reason of color-casting problem of underwater images. The underwater image colorfulness measure (UICM) can be defined as:

$$\text{UICM} = \lambda_1 \sqrt{\mu_{\alpha, RG}^2 + \mu_{\alpha, YB}^2} + \lambda_2 \sqrt{\sigma_{\alpha, RG}^2 + \sigma_{\alpha, YB}^2}, \quad (18)$$

where  $\lambda_1 = -0.0268$  and  $\lambda_2 = 0.1586$  [56],  $\mu_{\alpha, RG}$  and  $\mu_{\alpha, YB}$  are the  $\alpha$ -trimmed mean [57] of RG and YB color plane computed by the following equations:

$$\mu_{\alpha, RG} = \frac{1}{K - T_{\alpha R} - T_{\alpha L}} \sum_{i=T_{\alpha L}+1}^{K-T_{\alpha R}} \text{Intensity}_{RG,i}, \quad (19)$$

$$\mu_{\alpha, YB} = \frac{1}{K - T_{\alpha R} - T_{\alpha L}} \sum_{i=T_{\alpha L}+1}^{K-T_{\alpha R}} \text{Intensity}_{YB,i}, \quad (20)$$

and  $\sigma_{\alpha, RG}^2, \sigma_{\alpha, YB}^2$  are the corresponding second-order statistic  $\alpha$ -trimmed variance of RG and YB color plane:

$$\sigma_{\alpha, RG}^2 = \frac{1}{K} \sum_{p=1}^K (\text{Intensity}_{RG,p} - \mu_{\alpha, RG})^2, \quad (21)$$

$$\sigma_{\alpha, YB}^2 = \frac{1}{K} \sum_{p=1}^K (\text{Intensity}_{YB,p} - \mu_{\alpha, YB})^2. \quad (22)$$

With the complex underwater environment, the underwater images suffer serious blurriness due to the forward scattering of light by plankton, sand and minerals. The underwater image sharpness measure (UISM) can be calculated as:

$$\text{UISM} = \sum_{c=1}^3 \lambda_c \text{EME}(\text{grayscale edge}_c), \quad (23)$$

where  $c$  is the color channel of the RGM image,  $\lambda_c, c \in \{R, G, B\}$  indicate the relative visual responses of the RGB color components, and  $\text{EME}(\cdot)$  is the enhancement measure estimation function defined as:

$$\text{EME} = \frac{2}{k_1 k_2} \sum_{l=1}^{k_1} \sum_{k=1}^{k_2} \log\left(\frac{I_{\max,k,l}}{I_{\min,k,l}}\right), \quad (24)$$

where  $k_1 k_2$  is the number of image blocks,  $I_{\max,k,l}$  and  $I_{\min,k,l}$  are the local maximum and minimum in each block.

The stereoscopic disappearance of underwater images is caused by the contrast-degradation due to the backward scattering of light. The underwater image contrast measure (UIConM) is defined as the logAMEE operated on the grayscale  $I_{\text{gray}}$  image of the RGB image:

$$\text{UIConM} = \log\text{AMEE}(I_{\text{gray}}), \quad (25)$$



where  $I_{gray}$  is defined as  $(R + G + B)/3$  in this paper, and the  $\logAMEE(\cdot)$  is the log-Agaian measure of enhancement by entropy defined by the following equation[58]:

$$\logAMEE = \frac{1}{k_1 k_2} \otimes \sum_{l=1}^{k_1} \sum_{k=1}^{k_2} \frac{I_{max,k,l} \Theta I_{min,k,l}}{I_{max,k,l} \oplus I_{min,k,l}} \times \log\left(\frac{I_{max,k,l} \Theta I_{min,k,l}}{I_{max,k,l} \oplus I_{min,k,l}}\right), \quad (26)$$

where  $k_1 k_2$  is the number of image blocks, and  $\otimes$ ,  $\Theta$ ,  $\oplus$ ,  $\times$  denote the parameterized logarithmic image processing (PLIP) operations inspired by the human visual system: scalar multiplication, subtraction, addition, multiplication.

Based on the fundamental characteristics of the relationships between the above-mentioned features and the subjective evaluation, we build a non-linear regression model to correspond the features vector to the subjective MOS. Similar to the approaches used in [51], [59], we use Support Vector Regression (SVR) to estimate the image quality by the underwater enhancement features (UICM, UISM, and UICoM).

Let  $UIF_i = (UICM_i, UISM_i, UICoM_i)$  denotes the feature vector of the  $i$ -th image, and  $MOS_i$  is the MOS of the  $i$ -th image. The training data can be expressed as  $D_{train} = \{(UIF_i, MOS_i)\}_{i=1}^r$ . The standard SVR is modelled by the following optimal problem:

$$\begin{aligned} \min_{\mathbf{w}, b, \xi_i, \xi_i^*} \quad & \frac{1}{2} \|\mathbf{w}\|^2 + \lambda \sum_{i=1}^r (\xi_i + \xi_i^*) \\ \text{s.t.} \quad & \mathbf{w}^T \phi(UIF_i) + b - MOS_i \leq \epsilon + \xi_i, \\ & MOS_i - \mathbf{w}^T \phi(UIF_i) - b \leq \xi_i^*, \\ & \xi_i, \xi_i^* \geq 0, i = 1, \dots, r. \end{aligned} \quad (27)$$

where  $\phi$  is the low-to-high feature map,  $\lambda > 0$  is the model parameter, and  $\xi_i, \xi_i^*$  indicate the violated degree of each training sample. This SVR training settings were configured as follows: a) The radial basis function (RBF) was adopted as the kernel function to fit a non-linear map; b) 10-folder cross-validation approach was used to guarantee the generalized stability of the SVR model. Let  $\alpha_i, \alpha_i^*, i = 1, \dots, r$  be the solution of the dual problem of Equation (27) which formulated the final trained model denoted as UEIQM (underwater enhancement image quality metric).

## 2) Performance evaluation

As prescribed by the Video Quality Experts Group (VQEG) [60], we evaluate the performance of objective image quality metrics by quantifying their ability to predict subjective ratings (i.e., MOS) contained in our UEIQA database, using Pearson linear correlation coefficient (PLCC), Spearman rank order correlation coefficient (SROCC), Kendalls rank order correlation coefficient (KROCC) and root mean square error (RMSE). As suggested by VQEG [60], to account for any nonlinearity due to the subjective rating process and to facilitate comparison of metrics in a common analysis space, a nonlinear regression is first applied to the objective image quality score (IQS) and fitted to the [MOS, IQS], using the following function:

$$f(x_c) = \beta_1 \left\{ \frac{1}{2} - \frac{1}{1 + \exp[\beta_2(x_c - \beta_3)]} \right\} + \beta_4 x_c + \beta_5, \quad (28)$$

where  $x_c$  indicates the objective IQS, and  $\beta_i, i = 1, \dots, 5$  indicate the parameters for fitting of logistic regression. Once the nonlinear transformation was applied to the output of an objective metric, the PLCC, SROCC, KROCC and RMSE are computed between the subjectively measured MOS and the objective IQS.

Fig. 5 shows the scatter plots of the MOS versus BRISQUE, NIQE, BLIIND-II, UCIQE, UIQM and our proposed UEIQM, respectively. Table I lists the results of the PLCC, SROCC, KROCC and RMSE. The figure and table demonstrate that our proposed metric UEIQM outperforms other metrics in the prediction of the quality of enhanced underwater images. The good performance of the proposed metric demonstrates the importance of taking into account the specific characteristics of underwater environment. To verify whether the performance comparison, as shown in Table I, is statistically significant, hypothesis testing is conducted. As suggested in [60], the test is based on the residuals between the MOS and the quality predicted by a metric (i.e., referred to as M-MOS residuals). First, we evaluate the assumption of normality of the M-MOS residuals. When paired M-DMOS residuals are both normally distributed, an independent samples t-test is performed; otherwise, in the case of non-normality, a non-parametric version (i.e., Mann-Whitney U test) analogy to a t-test is conducted. The test results are given in Table II. This means the proposed metric is statistically significantly better than all other six state-of-the-art metrics.

TABLE I  
OVERALL PERFORMANCE COMPARISON OF NIQE, BRISQUE, BLIIND-II, UIQM, UCIQE AND OUR PROPOSED UEIQM

Metrics	PLCC	SROCC	KROCC	RMSE
BRISQUE	0.4148	0.4024	0.2737	1.9200
NIQE	0.3999	0.2366	0.1662	2.1998
BLIIND-II	0.1718	0.2708	0.1948	2.0779
UCIQE	0.5230	0.2886	0.1832	1.7977
UIQM	0.6159	0.5691	0.4101	1.6617
Proposed UEIQM	<b>0.8055</b>	<b>0.7657</b>	<b>0.5806</b>	<b>1.2508</b>

TABLE II  
RESULTS OF STATISTICAL SIGNIFICANCE TESTING BASED ON M-MOS RESIDUALS. 1 MEANS THAT THE DIFFERENCE (AS SHOWN IN TABLE I) IN PERFORMANCE IS STATISTICALLY SIGNIFICANT. 0 MEANS THAT THE DIFFERENCE (AS SHOWN IN TABLE I) IS NOT STATISTICALLY SIGNIFICANT

versus	BRISQUE	NIQE	BLIIND-II	UCIQE	UIQM
Proposed UEIQM	1 (Sig.)	1 (Sig.)	1 (Sig.)	1 (Sig.)	1 (Sig.)

## IV. CONCLUSION

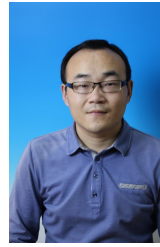
We have investigated five state-of-the-art enhancement algorithms and their impact on the perceptual quality of underwater images. Through human perception experiments, we found that DS algorithm degrades quality rather than enhancing the quality of images. Other four algorithms produce good results,

and IF and NLD significantly outperform other algorithms in enhancing image quality. This suggests that having a ground truth benchmark would be useful to faithfully compare enhancement algorithms and evaluate their capability and suitability for this problem. More enhancement algorithms are available in the literature or will be developed in the future. Therefore, constructing and sharing more subjective image quality databases is essential to facilitate research in this field. We have also indicated that a much faster approach to evaluating the output quality of enhancement algorithms is the use of an objective metric. Current metrics are unsuitable for application to the results of underwater image enhancement, and work is needed to devise new metrics. A proof of concept metric has been proposed, and there is still room for improvement. Future work will focus on developing a more sophisticated metric to approximate human judgment of the output quality of underwater image enhancement.

## REFERENCES

- [1] H. Liu, Q. Xu, S. Liu, L. Zhang, and H. Yang, "Evaluation of body weight of sea cucumber *apostichopus japonicus* by computer vision," *Chinese journal of oceanology and limnology*, vol. 33, no. 1, pp. 114–120, 2015.
- [2] Y. Y. Schechner and N. Karpel, "Recovery of underwater visibility and structure by polarization analysis," *IEEE Journal of oceanic engineering*, vol. 30, no. 3, pp. 570–587, 2005.
- [3] S. Q. Duntley, "Light in the sea," *JOSA*, vol. 53, no. 2, pp. 214–233, 1963.
- [4] V. Murino and A. Trucco, "Underwater computer vision and pattern recognition," *Computer Vision and Image Understanding*, vol. 79, no. 1, pp. 1–3, 2000.
- [5] D. Lee, S. Redd, R. Schoenberger, X. Xu, and P. Zhan, "An automated fish species classification and migration monitoring system," in *IECON'03. 29th Annual Conference of the IEEE Industrial Electronics Society (IEEE Cat. No. 03CH37468)*, vol. 2. IEEE, 2003, pp. 1080–1085.
- [6] S. G. Narasimhan and S. K. Nayar, "Chromatic framework for vision in bad weather," in *Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No. PR00662)*, vol. 1. IEEE, 2000, pp. 598–605.
- [7] —, "Vision and the atmosphere," *International journal of computer vision*, vol. 48, no. 3, pp. 233–254, 2002.
- [8] R. Fattal, "Single image dehazing," *ACM transactions on graphics (TOG)*, vol. 27, no. 3, p. 72, 2008.
- [9] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *IEEE transactions on pattern analysis and machine intelligence*, vol. 33, no. 12, pp. 2341–2353, 2010.
- [10] H.-Y. Yang, P.-Y. Chen, C.-C. Huang, Y.-Z. Zhuang, and Y.-H. Shiau, "Low complexity underwater image enhancement based on dark channel prior," in *2011 Second International Conference on Innovations in Bio-Inspired Computing and Applications*. IEEE, 2011, pp. 17–20.
- [11] C.-Y. Li, J.-C. Guo, R.-M. Cong, Y.-W. Pang, and B. Wang, "Underwater image enhancement by dehazing with minimum information loss and histogram distribution prior," *IEEE Transactions on Image Processing*, vol. 25, no. 12, pp. 5664–5677, 2016.
- [12] Y.-T. Peng, K. Cao, and P. C. Cosman, "Generalization of the dark channel prior for single image restoration," *IEEE Transactions on Image Processing*, vol. 27, no. 6, pp. 2856–2868, 2018.
- [13] D. Berman, S. Avidan *et al.*, "Non-local image dehazing," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 1674–1682.
- [14] Y.-T. Peng and P. C. Cosman, "Underwater image restoration based on image blurriness and light absorption," *IEEE transactions on image processing*, vol. 26, no. 4, pp. 1579–1594, 2017.
- [15] K. Zuiderveld, "Contrast limited adaptive histogram equalization," in *Graphics gems IV*. Academic Press Professional, Inc., 1994, pp. 474–485.
- [16] M. Zhao, C. Zhang, W. Zhang, W. Li, and J. Zhang, "Decorrelation-stretch based cloud detection for total sky images," in *2015 Visual Communications and Image Processing (VCIP)*. IEEE, 2015, pp. 1–4.
- [17] C. O. Ancuti and C. Ancuti, "Single image dehazing by multi-scale fusion," *IEEE Transactions on Image Processing*, vol. 22, no. 8, pp. 3271–3282, 2013.
- [18] J. Li, K. A. Skinner, R. M. Eustice, and M. Johnson-Roberson, "Watergan: Unsupervised generative network to enable real-time color correction of monocular underwater images," *IEEE Robotics and Automation letters*, vol. 3, no. 1, pp. 387–394, 2017.
- [19] Y. Wang, J. Zhang, Y. Cao, and Z. Wang, "A deep cnn method for underwater image enhancement," in *2017 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2017, pp. 1382–1386.
- [20] R. Hummel, "Image enhancement by histogram transformation," *Unknown*, 1975.
- [21] X. Qiao, J. Bao, H. Zhang, L. Zeng, and D. Li, "Underwater image quality enhancement of sea cucumbers based on improved histogram equalization and wavelet transform," *Information processing in agriculture*, vol. 4, no. 3, pp. 206–213, 2017.
- [22] A. S. A. Ghani, "Image contrast enhancement using an integration of recursive-overlapped contrast limited adaptive histogram specification and dual-image wavelet fusion for the high visibility of deep underwater image," *Ocean Engineering*, vol. 162, pp. 224–238, 2018.
- [23] C. Li and J. Guo, "Underwater image enhancement by dehazing and color correction," *Journal of Electronic Imaging*, vol. 24, no. 3, p. 033023, 2015.
- [24] C. Ancuti, C. O. Ancuti, T. Haber, and P. Bekaert, "Enhancing underwater images and videos by fusion," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2012, pp. 81–88.
- [25] J. Li, K. A. Skinner, R. M. Eustice, and M. Johnson-Roberson, "Watergan: Unsupervised generative network to enable real-time color correction of monocular underwater images," *IEEE Robotics and Automation Letters*, vol. 3, no. 1, pp. 387–394, 2018.
- [26] C. Fabbri, M. J. Islam, and J. Sattar, "Enhancing underwater imagery using generative adversarial networks," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*, 2018, pp. 7159–7165.
- [27] X. Ye, H. Xu, X. Ji, and R. Xu, "Underwater image enhancement using stacked generative adversarial networks," in *Advances in Multimedia Information Processing – PCM 2018*, R. Hong, W.-H. Cheng, T. Yamasaki, M. Wang, and C.-W. Ngo, Eds. Cham: Springer International Publishing, 2018, pp. 514–524.
- [28] C. Li, C. Guo, W. Ren, R. Cong, J. Hou, S. Kwong, and D. Tao, "An underwater image enhancement benchmark dataset and beyond," *IEEE Transactions on Image Processing*, vol. 29, pp. 4376–4389, 2020.
- [29] P. Guo, D. Zeng, Y. Tian, S. Liu, H. Liu, and D. Li, "Multi-scale enhancement fusion for underwater sea cucumber images based on human visual system modelling," *Computers and Electronics in Agriculture*, vol. 175, p. 105608, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0168169920304580>
- [30] A. Galdran, T. Araújo, A. M. Mendonça, and A. Campilho, "Retinal image quality assessment by mean-subtracted contrast-normalized coefficients," in *European Congress on Computational Methods in Applied Sciences and Engineering*. Springer, 2017, pp. 844–853.
- [31] D. L. Ruderman and W. Bialek, "Statistics of natural images: Scaling in the woods," in *Advances in neural information processing systems*, 1994, pp. 551–558.
- [32] H. R. Sheikh, M. F. Sabir, and A. C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," *IEEE Transactions on image processing*, vol. 15, no. 11, pp. 3440–3451, 2006.
- [33] R. I.-R. BT, "Methodology for the subjective assessment of the quality of television pictures," *International Telecommunication Union*, 2002.
- [34] J. Li, L. Krasula, Y. Baveye, Z. Li, and P. Le Callet, "Accann: A new subjective assessment methodology for measuring acceptability and annoyance of quality of experience," *IEEE Transactions on Multimedia*, vol. 21, no. 10, pp. 2589–2602, 2019.
- [35] L. Janowski and M. Pinson, "The accuracy of subjects in a quality experiment: A theoretical subject model," *IEEE Transactions on Multimedia*, vol. 17, no. 12, pp. 2210–2224, 2015.
- [36] T. Xiang, Y. Yang, and S. Guo, "Blind night-time image quality assessment: Subjective and objective approaches," *IEEE Transactions on Multimedia*, 2019.
- [37] Z. Wang and A. C. Bovik, "Mean squared error: Love it or leave it? a new look at signal fidelity measures," *IEEE signal processing magazine*, vol. 26, no. 1, pp. 98–117, 2009.
- [38] H. Hadizadeh and I. V. Bajić, "Full-reference objective quality assessment of tone-mapped images," *IEEE Transactions on Multimedia*, vol. 20, no. 2, pp. 392–404, 2017.
- [39] L. Ma, S. Li, F. Zhang, and K. N. Ngan, "Reduced-reference image quality assessment using reorganized dct-based image representation," *IEEE Transactions on Multimedia*, vol. 13, no. 4, pp. 824–829, 2011.

- [40] Z. Wan, K. Gu, and D. Zhao, "Reduced reference stereoscopic image quality assessment using sparse representation and natural scene statistics," *IEEE Transactions on Multimedia*, 2019.
- [41] W. Chen, K. Gu, T. Zhao, G. Jiang, and P. Le Callet, "Semi-reference sonar image quality assessment based on task and visual perception," *IEEE Transactions on Multimedia*, 2020.
- [42] L. Li, W. Lin, X. Wang, G. Yang, K. Bahrami, and A. C. Kot, "No-reference image blur assessment based on discrete orthogonal moments," *IEEE Transactions on Cybernetics*, vol. 46, no. 1, pp. 39–50, 2016.
- [43] K. Gu, S. Wang, G. Zhai, S. Ma, X. Yang, W. Lin, W. Zhang, and W. Gao, "Blind quality assessment of tone-mapped images via analysis of information, naturalness, and structure," *IEEE Transactions on Multimedia*, vol. 18, no. 3, pp. 432–443, 2016.
- [44] L. Li, W. Xia, W. Lin, Y. Fang, and S. Wang, "No-reference and robust image sharpness evaluation based on multiscale spatial and spectral features," *IEEE Transactions on Multimedia*, vol. 19, no. 5, pp. 1030–1040, 2017.
- [45] X. Min, K. Gu, G. Zhai, J. Liu, X. Yang, and C. W. Chen, "Blind quality assessment based on pseudo-reference image," *IEEE Transactions on Multimedia*, vol. 20, no. 8, pp. 2049–2062, 2017.
- [46] Y. Liu, K. Gu, S. Wang, D. Zhao, and W. Gao, "Blind quality assessment of camera images based on low-level and high-level statistical features," *IEEE Transactions on Multimedia*, vol. 21, no. 1, pp. 135–146, 2018.
- [47] L. Li, H. Zhu, S. Zhao, G. Ding, and W. Lin, "Personality-assisted multi-task learning for generic and personalized image aesthetics assessment," *IEEE Transactions on Image Processing*, vol. 29, pp. 3898–3910, 2020.
- [48] H. Zhu, L. Li, J. Wu, W. Dong, and G. Shi, "Metaiqa: Deep meta-learning for no-reference image quality assessment," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 14 131–14 140.
- [49] W. Lin and C.-C. J. Kuo, "Perceptual visual quality metrics: A survey," *Journal of visual communication and image representation*, vol. 22, no. 4, pp. 297–312, 2011.
- [50] D. Li, T. Jiang, W. Lin, and M. Jiang, "Which has better visual quality: The clear blue sky or a blurry animal?" *IEEE Transactions on Multimedia*, vol. 21, no. 5, pp. 1221–1234, 2018.
- [51] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Transactions on image processing*, vol. 21, no. 12, pp. 4695–4708, 2012.
- [52] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a completely blind image quality analyzer," *IEEE Signal Processing Letters*, vol. 20, no. 3, pp. 209–212, 2012.
- [53] A. K. Moorthy and A. C. Bovik, "Blind image quality assessment: From natural scene statistics to perceptual quality," *IEEE transactions on Image Processing*, vol. 20, no. 12, pp. 3350–3364, 2011.
- [54] M. Yang and A. Sowmya, "An underwater color image quality evaluation metric," *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 6062–6071, 2015.
- [55] K. Panetta, C. Gao, and S. Agaian, "Human-visual-system-inspired underwater image quality measures," *IEEE Journal of Oceanic Engineering*, vol. 41, no. 3, pp. 541–551, 2015.
- [56] W. K. Pratt, *Digital Image Processing: PIKS Inside*. J. Wiley & Sons, 2001.
- [57] J. Bednar and T. Watt, "Alpha-trimmed means and their relationship to median filters," *IEEE Transactions on acoustics, speech, and signal processing*, vol. 32, no. 1, pp. 145–153, 1984.
- [58] K. Panetta, S. Agaian, Y. Zhou, and E. J. Wharton, "Parameterized logarithmic framework for image enhancement," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 41, no. 2, pp. 460–473, 2010.
- [59] L. Liu, B. Liu, H. Huang, and A. C. Bovik, "No-reference image quality assessment based on spatial and spectral entropies," *Signal Processing: Image Communication*, vol. 29, no. 8, pp. 856–863, 2014.
- [60] J. Antkowiak, T. Jamal Baina, F. V. Baroncini, N. Chateau, F. FranceT-lecom, A. C. F. Pessoa, F. Stephanie Colonnese, I. L. Contin, J. Caviedes, and F. Philips, "Final report from the video quality experts group on the validation of objective models of video quality assessment march 2000," 2000.



**Pengfei Guo** received the Ph.D. degree from the South China University of Technology, Guangzhou, China, in 2015. He is currently an Associate Professor with the School of Computing Science, Zhongkai University of Agriculture and Engineering, Guangzhou, China. He has been the visiting scholar of Cardiff University. His research interests include computer vision and image quality assessment.



**Lang He** is an undergraduate student in the School of Computer Science and Engineering at Sun Yat-sen University. His research interests include computer vision and image quality assessment.



**Shuangyin Liu** received the Ph.D. degree from the College of Information and Electrical Engineering, China Agricultural University, in 2014. He is currently a Professor with the School of Information Science and Technology, Zhongkai University of Agriculture and Engineering. His current research interests include intelligent information system of agriculture, and computational intelligence.



chine learning and their applications in image processing, and data analysis.

**Delu Zeng** received his Ph.D. degree in electronic and information engineering from South China University of Technology, China, in 2009. He is now a full professor in the School of Mathematics in South China University of Technology, China. He has been the visiting scholar of Columbia University, University of Oulu, University of Waterloo. He has been focusing his research in applied mathematics and its interdisciplinary applications. His research interests include numerical calculations, applications of partial differential equations, optimizations, machine learning and their applications in image processing, and data analysis.



**Hantao Liu** received the Ph.D. degree from the Delft University of Technology, Delft, The Netherlands in 2011. He is currently an Associate Professor with the School of Computer Science and Informatics, Cardiff University, Cardiff, U.K. He is an Associate Editor of the *IEEE Transactions on Human-Machine Systems* and the *IEEE Transactions on Multimedia*.